Itanium-Based Solutions in the Financial Market

An IDC White Paper

Analysts: Kara M. Yokley, Christopher G. Willard, Ph.D. and Earl Joseph II, Ph.D.

Financial Services Market

The face of the financial services market[1] is changing. Contributing to the rapid expansion of software solutions and services for this sector are the emergence of Internet trading and personal banking, expectation of T+1 settlement[2] and decimalization,[3] and increased globalization. Any of these changes on its own could create a daunting amount of new financial information for end users and institutions to sift through as well as increased peaks in transaction demand, but their simultaneous materialization has the potential to lead to information overload.

With an abundance of financial data so readily available, customers have become more demanding about the content they receive and the speed at which they receive it. The process — from analyzing data and trends across the entire securities market, performing complex pricing analytics, routing and executing trades, to the clearing and settlement of transactions — should be nearly instantaneous. In the aftermath of the first commercial wave of Internet-based solutions, the dispersal of information is no longer the problem; making sense of the resultant deluge is. Financial services firms require an infrastructure that is able to keep pace with the growing requirements of customers, and in particular, an infrastructure that can perform floating-point calculations on large sets of data without sacrificing speed.

In 1999, sales of servers to banking, finance, and insurance companies accounted for $9.3 billion. Of that, more than 50% was from the sale of Unix-based servers, and only 12% was from Windows NT-based

1 According to IDC, this market is composed of banking, investing, and insurance companies. Among the common applications for this segment are risk analysis, economic modeling and forecasting, and financial analysis.

2 T+1 settlement would require settling of equity transactions within 24 hours, rather than the 72 hours currently enforced.

3 Decimalization, in the equity markets, is the move from fractional share prices to decimal intervals. Some expect trading volumes to increase fivefold as a result of decimalization because more granularity can be built into bids and asks.

IDC
Analyze the Future

5 Speen Street • Framingham, MA 01701 • Phone (508) 872-8200 • Fax (508) 935-4015

machines, with the balance running in proprietary operating environments. Revenues for the total workstation market (Unix + NT) in the economic and financial modeling segment were $755 million. In contrast to the server market, 72% of workstation revenues were from Windows NT-based systems.

The majority of computationally intensive tasks are still performed by Unix machines while servers based on Windows NT architectures are used primarily for less demanding collaborative workloads, including email, groupware, and file and print server applications. Among those workloads most likely to run using Unix machines are:

- Analytic computation — econometric modeling and risk management
- Decision support — data warehousing and data analysis/mining
- Infrastructure — networking, proxy/caching/security, and systems management
- Application development of financial software running on 64-bit-capable systems

Analytic computation such as credit risk management, and decision support such as pricing analytics, both require complicated matrix arithmetic and will benefit from increased floating-point performance and memory throughput. More real addressable memory will benefit each of these subsegments by taking away the need to perform unwieldy memory segmentation. In order to keep up with the expansion of the financial services market, companies in this space need hardware solutions that are high-performance, scalable, economical, and reliable.

Software companies, following the lead of their customers, are porting some of their more computationally demanding software to run on systems based on Intel's IA-64 architecture. SunGard and RiskMetrics are two examples of companies looking for scalable and inexpensive back-end solutions to power their software. SunGard, an end-to-end financial services solutions supplier, is currently the largest financial services software company, with approximately 70% of all NASDAQ trades being transacted on its systems. RiskMetrics is a younger company focused on institutional and retail risk management and on bringing financial services solutions to the individual investor.

The Itanium processor, with its floating-point performance and large memory addressability, presents Intel-based systems vendors with an opportunity to enter into subsegments of the financial services market, such as risk management. According to some financial institutions, they are transitioning to Intel architecture because it offers the

opportunity to standardize on a single family of processors across the entire enterprise — IA-32 for office productivity and Itanium and future processors for more computationally intensive tasks.

Currently, many financial services firms support multiple platforms and operating systems (OSs). As a result, software companies such as SunGard must offer products to run on these many configurations. Its customers need software and hardware to manage huge sets of data coming from diverse sources throughout the enterprise. SunGard supports more than 350 natively developed software products, mostly for Unix/RISC platforms. Its customers are showing interest in Intel's 64-bit solution, so it has begun to offer software solutions to take advantage of that interest.

SunGard is creating software that will tackle complex risk management problems and run on Intel-based servers. SunGard believes the Itanium-based systems provide hardware solutions that will offer improvements over current functionality. It is porting nearly a dozen of its software applications to run on the Itanium-based platform. One such product is Panorama, which addresses the needs of large financial institutions that have a global presence. While Panorama is currently tuned and optimized for IA-32/NT-based hardware, the company is developing an Itanium-based port for the most demanding applications.

Panorama and SunGard's other risk management software products are helping them to create Web-enabled globalized transactions with real-time settlement through their T+0.com products. Because risk management across entire enterprises requires numerous calculations from diverse sets of data, plenty of real addressable memory, floating-point performance, and scalability is required to make these products perform to their fullest potential.

RiskMetrics is tackling a different set of challenges. It is a fairly young company, borne of an internal need for substantive risk analysis at J.P. Morgan. Today, RiskMetrics is parlaying that risk management know-how by expanding into new markets with its RiskGrades product.

Since being spun off from J.P. Morgan in 1998, the company has expanded its addressable market. Taking advantage of trends in the world of individual trading, it is moving beyond its core market of desktop-centric software for institutional traders to now embracing the application service provider (ASP) model. Its ASP solutions enable companies to deliver institutional quality tools to the individual investor. To make the transition a success, RiskMetrics needed a robust back end that would be scalable, reliable, and affordable.

RiskGrades is a tool that can calculate the risk and run "what-if" scenarios for an individual portfolio at any given time. RiskMetrics has licensed its software to online money management firms and media portals, which in turn make it available to end users. Performance is of

the essence — to calculate the risk of portfolios of multiple individuals from across the Internet requires speed and high floating-point precision.

Although IA-32 is sufficient for RiskMetrics' needs right now, the company knows that as business expands to include more markets from around the world, so will its compute requirements. As more institutions offer RiskGrades to their millions of individual end users, the scale of the problem will increase and the hardware requirements will increase as well. RiskMetrics needs a high-end hardware solution flexible enough to ensure that it can scale its operations to handle a risk management problem of any size.

Itanium's Technical Computing Capabilities

In technical computing, once you get the answer, it is no longer interesting.
— Ancient Cybernetic Proverb

Complex applications in the financial services arena have historically placed some of the greatest performance demands on computer processors. In technical computing, the solution of one problem inevitably leads to a set of new and more complex problems (i.e., risk management). The complexity of the next generation of problems drives requirements for more powerful and complex tools — including computer processors. Thus, the terms technical computing and high-performance computing have become virtually synonymous.

Advanced financial applications, much like traditional technical applications, begin with mathematical models of physical phenomena, products, financial instruments, artistic images, and so on and then simulate the effects of different physical forces, market activities, perspectives, and so forth on the model in an effort to determine how the modeled object will behave in the real world.

This modeling and simulation process leads to a number of requirements for computer processor architectures in such areas as floating-point performance, memory performance, and support for large data sets. This section briefly reviews these requirements and provides an overview of Intel's strategy for meeting these requirements with its Itanium processor.

Floating-Point Performance

Initial implementations of Itanium will run at 3.2GFLOPS for 64-bit operations.

Floating-point performance equates to raw calculations speed — how many adds, multiplies, divides, and so forth, the system can perform in a given amount of time. At the top level, performance can be calculated as the number of floating-point functional units multiplied by the cycle time of the processor. On the Itanium this number is increased by functional units that can perform two operations in a single pass. The Itanium is configured with two floating-point functional units known as floating-point multiply add calculations (FMACs). These units can multiply two values and add that result to a third value.

(Multiply/add operations are at the heart of many technical calculations such as matrix multiplies.) The FMAC can also perform single multiplications or additions. Thus, an Itanium running at 800MHz can produce four floating-point results a cycle for a peak 64-bit performance rating of 3.2 billion floating-point operations per second (GFLOPS).

The Itanium architecture also includes two 32-bit FMACs that are tuned for 3D graphics performance and can each perform four floating-point operations per cycle for a 6.4GFLOPS single-precision rating on an 800MHz processor.[4]

It is important to note that these performance numbers will automatically increase with each step-up of clock rates in the IA-64 family. In addition, the IA-64 architecture is designed to allow future versions of the processor to be configured with additional FMACs.

The above analysis presents a best-case scenario in which the functional units are always busy. Although computer processors can maintain peak performance for only brief periods, Intel has incorporated a number of features in the Itanium architecture that help to maximize sustained performance. These include:

- **Pipelined functional units.** Arithmetic operations generally require more than one machine cycle to complete. A pipelining scheme is used to allow the FMACs to produce results each cycle. The arithmetic operations are broken into a set of independent steps, each requiring one machine cycle to complete. The FMACs perform arithmetic operations in an assembly-line fashion, with each step accepting data from the previous step and sending results to the next step. Thus, after the pipeline is full, a result is produced each cycle.

- **Dual-function arithmetic units.** A secondary benefit of the dual-function FMAC strategy is that the processor is able to use both functional units even when the distribution of adds and multiplies is biased toward one operation. For example, if a section of code performs only additions, both FMACs can be employed on the task. In contrast, a system with separate addition and multiplication functional units would use the adder but would have to leave the multiply unit idle.

- **Large register sets.** The Itanium is configured with 128 floating-point registers. The more data that is directly available to the FMACs, the less likely a functional unit will stall due to lack of data. In addition, the large register sets provide a buffer for the memory system to move data in and out of memory.

---

4 The advantage of double-precision or 64-bit operations over single-precision or 32-bit operations is that the former allow larger sets of calculations to be performed before accumulated round-off errors begin to affect the accuracy of final results. 64-bit systems produce full 64-bit results each cycles; a 32-bit system running at the same clock speed generally requires two cycles to produce the same 64-bit results.

- **Internal parallelism.** The Itanium can issue up to six instruction per cycles, combining any of four integer arithmetic/logical operations, two load/store operations, two floating-point operations, and three branch operations. The ability to execute multiple operations not only keeps as much of the processor working as possible but also allows for the pre-fetching of data from memory into registers and cache, thus minimizing processor stalls due to data unavailability. The processor also enables a load-double pair instruction to feed the processor with a balance of a memory operation per floating-point operation.
- **Compiler support for parallelism.** The Itanium was designed to allow for closer coordination between the processor and the compilers that generate the machine instructions for the processor. Three instructions are bundled along with a template field where the compiler can provide "hints" to the hardware on the interactions between the instructions. These hints are used by the processor to schedule instructions in real time, and for pre-fetching of data for future operations.

Memory Performance — Keeping the Processor Fed

A large fraction of technical applications are memory-bound rather than compute-bound — that is, the speed of the memory system ultimately determines the speed of the application. Ideally, the memory system will move data in and out of the processors fast enough to keep the floating-point functional units from stalling for lack of data.

Memory performance is measured in terms of both latency (i.e., how many cycles it takes to get data from memory to the processor and, in so doing, fill cache lines for subsequent data requests) and bandwidth (i.e., how many bytes of data can be moved in a cycle). Current systems architectures use memory hierarchies to address both latency and bandwidth issues. Hierarchies consist of a main memory and several layers of caches, and they trade off memory speed for size and cost. A small, fast cache is located "close" to the registers and functional units and can supply data at roughly the rate the processor calls for it. Data is staged through successive levels of cache, with each level holding more data and running somewhat slower until main memory is reached.

At the base of the Itanium hierarchy is main memory, which can vary in size and speed depending on individual system configurations and the system bus (or chipset) that connects the Itanium processors to memory and I/O subsystems. The processor can read or write 16 bytes of data to/from memory every bus cycle; thus, for a 133MHz bus, the memory bandwidth is 2.1GBps. From memory, data is transferred into a 4MB L3 (level 3) cache. The L3 cache communicates with both the 96KB L2 cache and the register file, moving data at 11.7GBps (roughly 16 bytes per 800MHz system clock) and with a 24 cycle latency for floating-point numbers. The L2 cache feeds data directly

into the floating-point registers at a rate of 32 bytes of data per clock tick and with a 9 clock latency.

Although the L1 cache is bypassed by floating-point data, it is worth noting that it is divided into a 16KB instruction cache — L1I — and a 16KB integer data cache — L1D. Both caches operate on a 2 clock latency and provide fast, localized access for integer instructions and data.

Support for Large Data Sets

The Itanium processor is Intel's first 64-bit architecture. As such, it opens up new opportunities for Intel-based systems in technical markets. Scientific and engineering problems grow in two directions. First, more advanced analysis of current problems requires more detailed models, thus requiring larger, more detailed data sets. Second, next generation problems tend to involve more complex product designs or attempt to model more complex phenomena, and thus require large data sets to describe the problems (e.g., to replace a wind tunnel requires the integration of computational fluid dynamics [CFD] airflow, thermal analysis, physics, and chemistry into a single analysis that is 1,000 times more complex than today's computers are capable of processing).

The requirements to operate on larger data sets generates in turn requirements for computer systems to provide larger real and virtual memories. A computer system's addressable memory is usually determined by the size of its integer or address registers. 32-bit architectures can directly address 4GB of either real or virtual memory. Beyond this limit, some form of memory segmentation must be employed.

64-bit architectures can in theory address about 1019 (10^19) bytes of data, a number that is so large that it defeats the authors' ability to describe it in any meaningful terms. The important point is that 64-bit systems allow computer systems to expand memory virtually indefinitely without having to resort to some form of segmentation. This large memory space has two major advantages for technical users:

1. **Increased applications performance.** A major bottleneck for many technical applications is the time spent in swapping data between disk and memory — access to data on disk is roughly an order of magnitude slower than memory. The large memories provided by 64-bit systems — up to 16GB on the Itanium — allow applications to keep more of the problem set in memory, thus reducing the amount of time spent reading and writing disk files. In the best case, the entire problem can be moved into memory, transforming an "out of core" problem into an "in core" problem.

2. **Simplified programming model.** The larger real and virtual memories afforded by 64-bit architectures enable applications developers to design programs without having to divide the problem into smaller, memory-sized segments and then develop a code to manage those segments.

IDC Analysis

Banks, credit houses, and online trading companies are already showing interest in the price/performance Itanium promises. However, transitions to new architectures rarely happen overnight. Itanium is just Intel's first iteration of its EPIC architecture, but it will have benefits inherent to high-end processors: increased addressable memory, improved floating-point performance, and better memory bandwidth. These benefits, in addition to Itanium's backwards compatibility with existing IA-32 processors, will give end users another option for 64-bit capable servers and will enable Intel to enter new workload segments and increase its presence in the financial services market.

NORTH AMERICA

Corporate Headquarters
5 Speen Street
Framingham, MA 01701
508-872-8200

IDC Canada
36 Toronto Street, Suite 950
Toronto, Ontario, Canada M5C2C5
416-369-0033

IDC Irvine
18831 Von Karmen Ave, Ste 200
Irvine, CA 92612
949-250-1960

IDC Mtn. View
2131 Landings Drive
Mountain View, CA 94043
650-691-0500

IDC New Jersey
120 Wood Ave South, Suite 509
Iselin, NJ 08830
732-632-9222

IDC New York
2 Park Avenue
Suite 1505
New York, NY 10016
212-726-0900

IDC Texas
100 Congress Ave, Suite 2000
Austin, TX 78701
512-469-6333

IDC Washington
8304 Professional Hill Drive
Fairfax, VA 22031
703-280-5161

ASIA/PACIFIC

IDC Asia/Pacific (Hong Kong)
12/Floor, St. John's Building
33 Garden Road
Central, Hong Kong
852-2530-3831

IDC Asia/Pacific (Singapore)
71 Bencoolen Street, #02-01
Singapore 189643
65-226-0330

IDC Australia
Level 4, 76 Berry Street
North Sydney, NSW 2060, Australia
61-2-9922-5300

IDC China
Room 611, Beijing Times Square,
88 West Chang'an Avenue, Beijing,
P.R. China, 100031
86-10-8391-3456

IDC (India) Limited
Cyber House
B-35, Sector 32 - Institutional
Gurgaon - 122002, Haryana, India
91-124-6381673 to 80

IDC Japan
10F The Itoyama Tower
3-7-18, Mita Minato-ku
Tokyo 108-0073, Japan
81-3-5440-3400

IDC Korea Ltd
Suite 704, Korea Trade Center
159-1, Samsung-Dong, Kangnam-Ku
Seoul, Korea 135-729
82-2-55-14380

IDC Malaysia
Suite 13-03, Level 13, Wisma KiaPeng
No. 3, Jalan Kia Peng
50450 Kuala Lumpur, Malaysia
6-03-2163 3715

IDC New Zealand
Level 7, 246 Queen Street
Auckland, New Zealand
64-9-309-8252

IDC Philippines
7F, SEDCCO 1Bldg
Rada Street Corner
Legaspi Street
Legaspi Village
Makati City, Philippines
632-894-4808

IDC Taiwan Ltd.
10F, 31
Jen-Ai Rd, Sec 4,
Taipei 106, Taiwan, R.O.C.
886-2-2731-7288

IDC Thailand
27 Soi Charoen Nakorn 14
Charoen Nakorn Road, Klongtonsai
Klongsan Bangkok 10600, Thailand
66-2-439-4591-2

IDC Vietnam
37 Ton Duc Thang Street
Unit 1606
District-1 Hochiminh City Vietnam
84-8-910-1235

EUROPE, MIDDLE EAST, AND AFRICA

IDC Austria
c/o Loisel, Spiel, Zach Consulting
Mayerhofgasse 6
A-1040 Vienna, Austria
43-1-50-50-900

IDC Benelux (Belgium)
29 Avenue Louis Gribaumont
B-1150 Brussels, Belgium
32-2-779-46-04

IDC Benelux (The Netherlands)
A. Fokkerweg 1
1059 CM Amsterdam
The Netherlands
31-20-669-2721

IDC Central Europe (ECE)
Male Namesti 13
Praha 1 110 00, Czech Republic
420-2-2142-3140

IDC Central Europe (Germany)
Nibelungenplatz 3, 11th Floor
60318 Frankfurt, Germany
49-69-90502-0

IDC Central Europe (Switzerland)
Niederlassung Züerich
WTC, Leutschenbachstrasse 95
CH - 8050 Züerich
Switzerland
41-1-307-1000

IDC Egypt
39 Iraq Street
Mohandesseen, Cairo, Egypt
20-2-336-7355

IDC France
Immeuble La Fayette
2, Place des Vosges, Cedex 65
92051 Paris la Defense 5, France
33-14-904-8000

IDC Hungary
Bajcsy-Zsilinszky út. 57
Building 3, Rooms 103-104
H-1065 Budapest, Hungary
36-1-153-0555/ext. 165, 166

IDC Israel
4 Gershon Street
Tel Aviv 67017, Israel

IDC Italy
Viale Monza, 14
20127 Milano, Italy
390-2-284-571

IDC Nigeria
House 2, 'C' Close
403 Road, 4th Avenue
New Extension, Festac Town
Lagos, Nigeria
234-1-883585

IDC Nordic (Denmark)
Jagtvej 169B
DK-2100 Copenhagen, Denmark
45-39-162222

IDC Nordic (Finland)
Jarrumiehenkatu 2
FIN-00520
Helsinki, Finland
358-9-8770-466

IDC Nordic (Sweden)
Box 1096 Kistagången 21
S-164 25 Kista, Sweden
46-8-751-0415

IDC Poland/ProMarket
Wrobla 43
02-736 Warsaw, Poland
48-22-644-4105

IDC Portugal
Av. Antonio Serpa, 36 Piso 9
1050-027 Lisbon
Portugal
351-21-796-5487

IDC Russia
c/o PX Post, RDS 186
Ulitsa Zorge 10
Moscow 125525
Russian Federation
7-501-929-9959

IDC South Africa
c/o BMI-TechKnowledge
3rd Floor, 356 Rivonia Blvd.
PO Box 4603, Rivonia, 2128
South Africa
27-11-803-6412

IDC Spain
Ochandiano, 6
Centro Empresarial El Plantio
28023 Madrid
34-91-7080007

IDC Turkey
Tevfik Erdonmez Sok. 2/1 Gul Apt.
Kat 9D; 46 Esentepe
Istanbul, Turkey
90-212-275-0995

IDC U.K.
British Standards House
389 Chiswick High Road
London W4 4AE
United Kingdom
44-20-8987-7100

LATIN AMERICA

IDC Miami
Latin America Headquarters
5301 Blue Lagoon Drive
Suite 490
Miami, FL 33126
305-267-2616

IDC Argentina
Trends Consulting
Rivadavia 413, 4th Floor, Suite 6
C1002AAC, Buenos Aires, Argentina
54-11-4343-8899

IDC Brasil
Alameda Ribeirão Preto, 130 cj 41
01331-000 São Paulo
SP Brazil
55-11-253-7869

International Data Corp. Chile
Luis Thayer Ojeda 166 Piso 12
Providencia, Santiago 9, Chile
56-2-231-0111

IDC Colombia
Carrera 40 # 103-78
Bogota, Colombia
571-533-2326

IDC Mexico
Select - IDC
Av. Nuevo Leon No. 54 Desp. 501
Col. Hipodromo, Condesa
C.P. 06100 Mexico, D.F.
52-5-256-1426

IDC Venezuela
Calle Guaicapuro
Edif. Torre Seguros Alianza
Piso 6, Ofc. 6-D, El Rosal
Caracas 1060, Venezuela
58-2-951-3270

IDC delivers accurate, relevant, and high-impact data and insight on information technology to help organizations make sound business and technology decisions. IDC forecasts worldwide IT markets and adoption and technology trends, and analyzes IT products and vendors, using a combination of rigorous primary research and in-depth competitive analysis. IDC is committed to providing global research with local content through more than 500 analysts in more than 40 countries worldwide. IDC's customers comprise the world's leading IT suppliers, IT organizations, and the financial community. Additional information on IDC can be found on its Web site at http://www.idc.com.

IDC is a division of IDG, the world's leading IT media, research, and exposition company.

IDC
Analyze the Future